

AD-A249 333



STIC
ELECTR
APR 28 1992

(2)

The Counselor Project

Department of Computer and Information Science
University of Massachusetts
Amherst, Mass. 01003

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

Case-Based Reasoning in a
Rule-Governed Domain

Edwina L. Rissland
and
David B. Skalak

1988
CPTM #26

92-10698

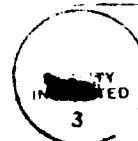


182-10698

92 4 10 182

This series of internal memos describes research in
artificial intelligence conducted under

Accession For	
DTIC Tab	<input checked="" type="checkbox"/>
Unpublished	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



**Case-Based Reasoning in a
Rule-Governed Domain**

**Edwina L. Rissland
and
David B. Skalak**

**1988
CPTM #26**

92 4 24 129

Case-Based Reasoning in a Rule-Governed Domain¹

Edwina L. Rissland and David B. Skalak
Department of Computer and Information Science
University of Massachusetts
Amherst, Massachusetts 01003
(413) 545-0332, RISSLAND@umass.cs.edu

Abstract

In this paper, we discuss the problem of reasoning with cases in a domain governed by rules, in particular, the problems of interpreting the meaning of terms (statutory predicates) used in the rules and combining case-based reasoning ("CBR") with other modes of reasoning, such as rule-based ("RBR") and model-based reasoning. Terms used in statutes are typically underdefined in the statute and inherently open-textured and thus require precedent-based reasoning to interpret their meaning for particular fact situations. In such domains one needs to combine reasoning about the rules (statutes) with the precedents that concern them.

We describe our precedent-based case-based reasoner TAX-HYPO that operates in the statutory domain of tax law. TAX-HYPO is a derivative system of our earlier CBR system HYPO which operated in the common law domain of trade secret law. We describe our system CABARET which is an environment to support (1) building precedent-based CBR systems like HYPO and TAX-HYPO; and (2) experimentation with mixed paradigm systems involving CBR. In particular, we discuss some of the heuristics we use to control and combine reasoning with cases and rules.

TOPIC: Both a principles and a case-study paper: After introduction to the topic, we present a "case-study" for TAX-HYPO in Section 2, and a discussion of "principles" and our general CBR tool CABARET in Section 3. **SUBTOPIC:** Case-Based Reasoning.

STATUS: One version of TAX-HYPO is completed. The CABARET environment is being researched. A second version of TAX-HYPO is being re-implemented in CABARET.

DOMAIN: Statutory law, in particular, tax law.

LANGUAGE: Common LISP.

EFFORT: TAX-HYPO: 1.5 person-years. CABARET: 1 person-year.

¹This work was supported (in part) by the Advanced Research Projects Agency of the Department of Defense, monitored by the Office of Naval Research under contract no. N00014-87-K-0238, and a grant from GTE Laboratories, Inc., Waltham, Mass.

Case-Based Reasoning in a Rule-Governed Domain*

Edwina L. Rissland and David B. Skalak[†]
Department of Computer and Information Science
University of Massachusetts
Amherst, Massachusetts 01003

September 19, 1988

Abstract

In this paper, we discuss the problem of reasoning with cases in a domain governed by rules, in particular, the problems of interpreting the meaning of terms (statutory predicates) used in the rules and combining case-based reasoning ("CBR") with other modes of reasoning, such as rule-based ("RBR") and model-based reasoning. Terms used in statutes are typically underdefined in the statute and inherently open-textured and thus require precedent-based reasoning to interpret their meaning for particular fact situations. In such domains one needs to combine reasoning about the rules (statutes) with the precedents that concern them.

We describe our precedent-based case-based reasoner TAX-HYPO that operates in the statutory domain of tax law. TAX-HYPO is a derivative system of our earlier CBR system HYPO which operated in the common law domain of trade secret law. We describe our system CABARET which is an environment to support (1) building precedent-based CBR systems like HYPO and TAX-HYPO; and (2) experimentation with mixed paradigm systems involving CBR. In particular, we discuss some of the heuristics we use to control and combine reasoning with cases and rules.

1. Introduction

Many domains of interest to AI have the following characteristics:

- There are rules, policies or codes setting forth a controlling model of the domain or action in it. That is, the domain is *rule-governed*.

*This work was supported (in part) by the Advanced Research Projects Agency of the Department of Defense, monitored by the Office of Naval Research under contract no. N00014-87-K-0238, and a grant from GTE Laboratories, Inc., Waltham, Mass.

[†]Copyright ©1988. Edwina L. Rissland and David B. Skalak. All rights reserved.

- These rules use terms whose meaning is unclear, for instance, because the term is not adequately defined in the rule set or because it can't be or is not intended to be. That is, the terms are underdefined or *open-textured*.
- There is a body of cases about the past application of the rules and interpretation of the terms used in them. That is, there are *cases*.
- The ambiguous terms must be interpreted by appeal to precedent cases addressing past interpretations. That is, the reasoning is *precedent-based*.
- A new fact situation must be analyzed by appeal to both the rules and the cases. That is, the reasoning involves *mixed paradigms*.

Tax law is a typical example of such a domain. So is the law involving the administration of such things as Social Security, housing regulations, and employee benefits. Non-legal domains sharing these characteristics are tactical planning, where there are rules governing courses of action as well as past episodes of action, and certain fields of design, where there are design methods and constraints as well as past designs. While mathematics has both rules (i.e., definitions and theorems) as well as cases (i.e., examples and counter-examples), the terms used are as far as one can get from open-textured since they are tightly defined. (Of course, in a nascent or rapidly changing mathematical area, things are not as "close-textured" as one might hope [Lakatos, 1976].) Thus in pure mathematical domains the problems of interpretation of interest to us here are minimal although the problem of mixed paradigm reasoning is still central.

The problem of open-textured terms ("predicates") is central in legal reasoning, especially in statutory domains. Open-textured statutory predicates often present hard problems of statutory interpretation -- hard for both attorneys and computer programs purporting to do legal reasoning.

Open-textured predicates can defeat unimaginative rule-based schemes. For instance, an inflexible forward-chaining production system would attempt to set up the antecedents of a rule as sufficient conditions to the firing of the conclusion. With an open-textured predicate in a conclusion, however, no fixed set of antecedents can provide sufficient conditions in all circumstances. There is no collection of conditions that one could provide, for example, as to whether a defendant had taken *reasonable care* to avoid an accident that caused harm to a tort plaintiff.¹ Usually, the most that a judge or a legal commentator can do is to set forth the factors that should be used to determine whether a condition like reasonable care is met.

Since Anglo-American law operates under the doctrine of precedent, *stare decisis*, the approach to the interpretation problem is of necessity case-based. A legal expert uses both rules and cases, sometimes working from one, sometimes from the other. The reasoning is often guided by the

¹This determination is critical in a tort action, since the exercise of reasonable care by the defendant will often secure a defendant from payment of damages to a plaintiff. However determining what constitutes reasonable care and whether it has been exercised in a particular situation is no little problem and traditional sources like dictionaries, treatises, commentator services are not much help. For instance, Black's Law Dictionary defines *reasonable care* as "That degree of care which a person of ordinary prudence would exercise in the same or similar circumstances. ... Due care under all the circumstances." [Black, 1979]

requirements of the rules or tied together by appeal to them and the arguments concerning interpretation are essentially precedential in character. Thus reasoning with cases in a rule-governed domain raises interesting problems of control: when to do what, how to combine and interleave two distinct styles of reasoning, how to share results across the reasoning modes, etc.

Although the law is a domain *par excellence* for studying problems of interpretation, open-textured predicates, combining CBR and RBR, etc., such questions also are present in domains currently treated largely by expert system ("ES") techniques (e.g., medicine). In such domains, ES approaches often either simply throw away a vast repository of expertise, namely cases or episodes of past reasoning, or pretend that the reasoning is solely rule-based when it is not. Improving performance of such ES applications will require something else — we hypothesize CBR. Consideration of CBR in concert with other reasoning paradigms also brings one closer to considering problems of learning and knowledge acquisition, which ultimately one cannot afford to ignore.

The problem of open-textured predicates, particularly their relation to the philosophical idea of the hard/easy case distinction, was investigated by Gardner [Gardner, 1987]. For Gardner, cases were used either as a check on rule-based reasoning or as a last resort when rule-based reasoning failed; thus, the role of CBR was somewhat subjugated to RBR. The problem of reasoning with precedents was attacked by Ashley [Ashley, 1988a] and our group and work reported on here builds on such efforts.

2. TAX-HYPO

2.1 When to Take a Deduction for Your Home Office

TAX-HYPO² takes its basic architecture from HYPO [Ashley, 1988a; Ashley and Rissland, 1988; Rissland and Ashley, 1987; Rissland and Ashley, 1986]: a case-based legal reasoning system composed of a library of legal cases, a library of indices referred to as "dimensions", and modules to retrieve relevant cases from its knowledge base, generate a graph of cases relevant to a given fact situation, create 3-ply legal arguments based upon the retrieved cases and their similarity along dimensions, and manufacture hypotheticals that test and challenge the arguments generated. The difference between HYPO and TAX-HYPO lies primarily in the type of legal domain that each addresses. HYPO was designed to work in the domain of trade secrets law, which is substantially a "common law" domain. A "common law" domain is governed by the law made by judges in the process of deciding litigated cases. TAX-HYPO works in a "statutory law" domain — one governed by law explicitly set down in legal statutes that have been promulgated by a legislature. TAX-HYPO's statutory area is a prototypical one: Federal income tax law, which is governed by the Internal Revenue Code (the "Code"). In particular, TAX-HYPO reasons with cases dealing with the so-called "home office deduction," governed by §280A(c)(1) of the Code:

²The title of this subsection notwithstanding, nothing in this paper is to be construed as legal advice, for which the reader should consult his or her own tax adviser.

[A deduction may be taken for any] item to the extent such item is allocable to a portion of the dwelling unit which is **EXCLUSIVELY USED** on a **REGULAR** basis --

(A) [as] the **PRINCIPAL PLACE OF BUSINESS** for any trade or business of the taxpayer,

(B) as a place of business which is used by patients, clients, or customers in **MEETING OR DEALING** with the taxpayer in the normal course of his trade or business, or

(C) in the case of a **SEPARATE STRUCTURE** which is not attached to the dwelling unit, in connection with the taxpayer's trade or business.

In the case of an employee, the preceding sentence shall apply only if the exclusive use referred to in the preceding sentence is for the **CONVENIENCE OF HIS EMPLOYER**. [I.R.C. §280A(c)(1), *capitalization supplied*.]

The "home office deduction" obviously deals with the circumstances under which a taxpayer may legitimately deduct on his Federal income tax return expenses relating to an office he maintains at his residence. A example situation from TAX-HYPO's Case Knowledge Base ("CKB") that brings this Code Section to bear is that of Max Frankel, Editor of *The New York Times*.³ Mr. Frankel maintained an office at his home in the Bronx, which he used for reading the morning papers, writing memoranda, clipping materials, and speaking by telephone to his employees, prominent politicians and community leaders. Although the Tax Court granted a home office deduction on grounds relating to a consulting position Mrs. Frankel had, it denied that Mr. Frankel met any of the three disjunctive requirements of the statute, (A), (B), or (C) above. In particular, the use of the telephone to conduct business was held not to satisfy the **MEETING OR DEALING** predicate, which was construed to require the physical presence of the business contacts.

The terms capitalized in the quoted statute actually appear in lower case in the Code. Here, these terms have been emphasized to highlight their role in TAX-HYPO as *statutory predicates*, important terms or phrases on which the meaning of the statute turns⁴. These are the phrases whose meaning taxpayers and the IRS often argue over in tax litigation. They provide indices into the research materials used by tax attorneys to resolve taxpayer inquiries. [CCH, 1988; Prentice Hall, 1988; BNA, 1986]. While the meaning of phrases used in the Code is sometimes partly clarified by official regulations thereunder issued by the Internal Revenue Service, statutory predicates are inherently open-textured. [Hart, 1961] The reach of their meaning is fundamentally unclear, varies greatly according to the factual context in which they are used, and defeats precise definition by rules. For clues as to the interpretation and scope of statutory predicates, TAX-HYPO and legal researchers rely on previously litigated cases that have construed these terms.

³Max and Tobia Frankel v. Commissioner, 82 USTC 318 (Filed February 28, 1984)

⁴Statutory predicates will be capitalized throughout.

For example, in *Meiers v. Comm.*⁵, a home office deduction case dealing with, among other things, the interpretation of PRINCIPAL PLACE OF BUSINESS, a Court of Appeals said

In determining the taxpayer's principal place of business, we think a major consideration ought to be the length of time the taxpayer spends in the home office as opposed to other locations. . . . There are other factors, which may from time to time weigh in the balance, such as the importance of the business functions performed by the taxpayer in the home office; the business necessity of maintaining a home office; and the expenditures of the taxpayer to establish a home office.⁶

This is precisely the sort of problem for which the precedent-based CBR methods of HYPO and TAX-HYPO are ideally suited.

2.2 Knowledge Representation in TAX-HYPO: Statutory Predicates, Dimensions and Cases

Unlike HYPO, which did not involve the analysis of a statute, TAX-HYPO associates with each statutory predicate in §280A(c)(1)

- EXCLUSIVE USE
- REGULAR USE
- PRINCIPAL PLACE OF BUSINESS
- MEETING OR DEALING
- SEPARATE STRUCTURE and
- CONVENIENCE OF EMPLOYER

a subset of the 15 dimensions currently implemented in its index library. In TAX-HYPO, dimensions capture the substantive legal features of a fact situation used to determine whether a statutory predicate is satisfied.⁷

For example, indexed under the predicate EXCLUSIVE USE, §280A's overarching requirement that the home office be used exclusively for business activities, are the implemented dimensions:

- *use for any personal purposes* (recording whether any personal, even *de minimis* activities undertaken in the home office, such as personal bill paying, personal telephoning and television viewing),

⁵ 782 F.2d 75 (7th Cir. 1986) (home office deduction granted to couple managing a laundromat, who used a home office to draft work schedules and perform bookkeeping.)

⁶ *Ibid.*, at 79.

⁷ Dimensions were identified by resident attorneys and researchers working in the Case-Based Reasoning Group at UMass, based on a close reading of materials used by tax practitioners and scholars, as well as of cases deciding issues under §280A and related Code Sections.

- *physical separation within unit* (capturing whether the space claimed as the home office is distinct from other, personal spaces), and
- *presence of personal furnishings* (referring to the mainly circumstantial evidence that non-business effects are present in the home office, such as televisions or houseguest accommodations).

Having more or less of each such dimension clearly influences one's strength or weakness with regard to EXCLUSIVE USE.

Using the machinery of dimensional analysis, TAX-HYPO dynamically applies its library of dimensions as indices into lines of cases in the CKB that construe one or more of the statutory predicates. Some dimensions also provide an indication of how strongly a feature of a case argues for the satisfaction of a statutory predicate. For example, the more personal furnishings present, the stronger the case for the IRS along the dimension *presence of personal furnishings*. Obviously, the dimensions can also act as indices into the case base. Cases indexed under the dimension *presence of personal furnishings*, for example, include CKB cases Thalacker⁸, Chauls⁹, Gomez¹⁰, and Weightman¹¹.

The underlying idea is that the connection between the open-textured phrases in the statute, whose meaning we are trying to determine, and the cases that construe those phrases is made by interposing dimensions. Dimensions are indexed by predicates; in turn, dimensions index cases.

TAX-HYPO currently has a CKB of approximately 20 litigated, reported tax cases and a handful of hypothetical cases, all of which are stored as a hierarchical collection of frames. Frames represent aspects of a tax case such as the taxpayer, his employer, the nature of his job, and features of his home office. The six statutory predicates are indexed by the 15 dimensions that have been implemented. Approximately 15 dimensions have been identified but not implemented.

2.3 An Example of TAX-HYPO at Work

Consider the problem of a hypothetical taxpayer, Ms. Olivia, who wants to determine whether she meets the statutory requirements for a home-office deduction. Ms. Olivia is a concert musician with the Metropolitan Opera Orchestra. She is not provided with a practice room by the Opera Association, and so practices over 30 hours a week in a dedicated practice studio in her New York City apartment. Since she does not meet or deal with clients in the home office (disjunct (B) of §280A(c)(1)), and the practice studio is not in a separate structure (disjunct C), she must show that her Upper West Side apartment is her principal place of business (disjunct A). Indexed under the statutory predicate PRINCIPAL PLACE OF BUSINESS are the dimensions

⁸Jimmy Thalacker v. Commissioner, 48 T.C.M. 1104 (Filed September 11, 1984)

⁹Robert Chauls v. Commissioner, 51 T.C.M. 234 (Filed October 22, 1980)

¹⁰Sharon L. Gomez v. Commissioner, 41 T.C.M. 585 (Filed December 18, 1980)

¹¹George H. Weightman v. Commissioner, 42 T.C.M. 104 (Filed June 18, 1981)

- *income from home office* (percent of income derived from activities undertaken in home office),
- *primary responsibility location* (where the taxpayer's primary business responsibilities are discharged) and
- *relative time in home office* (the percent of working hours the taxpayer spends in his home office).

When presented with an encoding of Ms. Olivia's tax situation in TAX-HYPO's Case Representation Language, TAX-HYPO retrieved five relevant cases from its CKB. Each provides an interpretation of the phrase PRINCIPAL PLACE OF BUSINESS: Drucker, Meiers, Weissman, Honan, and Baie. An excerpt of a dump of the Case Analysis Record for Ms. Olivia's case provides, in fact:

#<ACR : |MS-OLIVIA V. COMM. : ACR|>, an object of flavor ACR,
has instance variable values:

NAME: |MS-OLIVIA V. COMM. : ACR|
CASE: #<CASE : |Ms-Olivia v. Comm.|>

APPLICABLE-DIMENSIONS:

(#<DIMENSION : PRIMARY-RESPONSIBILITY-LOCATION>
#<DIMENSION : RELATIVE-TIME-IN-HOME-OFFICE>
#<DIMENSION : EXISTS-FREQUENCY-OF-USE>
#<DIMENSION : SPECIFIC-FREQUENCY-OF-USE>
#<DIMENSION : IN-SEPARATE-STRUCTURE>
#<DIMENSION : NECESSARY-TO-PERFORM-DUTIES>
#<DIMENSION : CONDITION-OF-EMPLOYMENT>)

NEAR-MISS-DIMENSIONS:

(#<DIMENSION : EXISTS-EXPENSE-TO-ESTABLISH>
#<DIMENSION : EXISTS-PHYSICAL-CONTACT>
#<DIMENSION : PHYSICAL-SEPARATION-WITHIN-UNIT>
#<DIMENSION : PRESENCE-OF-PERSONAL-FURNISHINGS>
#<DIMENSION : USE-FOR-ANY-PERSONAL-PURPOSES>)

MOST-ON-POINT-CASES-BY-CLAIM:

((#<TYPE-OF-CLAIM : PRINCIPAL-PLACE-OF-BUSINESS>
(PLAINTIFF #<CASE : |Weissman v. Comm.|>
#<CASE : |Meiers v. Comm.|>

```

#<CASE : |Drucker v. Comm.|>)
(DEFENDANT #<CASE : |Baie v. Comm.|>
#<CASE : |Honan v. Comm.|>))
(<Most on point cases for other claims ...>)

```

TAX-HYPO uses the Case Analysis Record at this point to generate a claim lattice of retrieved cases, as well as 3-ply arguments and hypotheticals addressing the interpretation of the statutory predicate PRINCIPAL PLACE OF BUSINESS for the case of Ms. Olivia. By using TAX-HYPO on each statutory predicate one is able to construct arguments which could then be combined to address Ms. Olivia's home office deduction. Thus, by using TAX-HYPO "in parallel" with itself one can deal with a case involving a statutory rule.

2.4 Evaluation of TAX-HYPO

TAX-HYPO was partially conceived as and proved to be a vehicle to demonstrate the viability of the HYPO architecture in a different kind of legal domain. Specifically, TAX-HYPO provided a second test for the efficacy of using dimensions as indices that capture substantive domain knowledge. In particular, dimensions were used successfully to indicate the strength, or weakness, of a litigant's position along relevant features, without resort to numerical artifice. Using the argument-generating modules of HYPO, 3-ply legal arguments were also successfully generated in the Federal income tax domain, which cited supporting cases for one side, distinguished those supporting cases on behalf of the opposition and cited opposing counter-examples, and provided opportunity for the first side to distinguish any cited counterexamples.

However, the experience of applying HYPO's methodology to a rule-governed area of the law exposed a need for additional functionality not provided by TAX-HYPO. The user of TAX-HYPO can perform dimension analysis, display claim lattices, generate arguments, etc., only with respect to one statutory predicate, and display simultaneously the results of its analysis of several predicates. But what TAX-HYPO lacks is the capacity to combine the analyses of individual predicates to generate an argument that takes into account the statute as a whole. The program fails to argue or evaluate the claim for a deduction: it can merely argue the pros and cons of specific individual predicates. Particularly, TAX-HYPO is not cognizant of the Boolean connectives in the statute, and treats it merely as an set of unrelated predicates. The current implementation has no heuristics for determining the *focus of attention* of its analysis and is not at all smart about which statutory predicates to go after in a given fact situation.

Additionally, the home office deduction statute is informed by regulations and lives within a complex network of other Federal and state tax statutes. These regulations and other statutes can be expressed as rules, with their own open-textured predicates that require case-based treatment.

So TAX-HYPO fails to take sufficient account of the rule-like aspects of the statute at issue, fails to take account of related rules, and fails to integrate deductive reasoning with such rules with an inductive, case-based approach. For these reasons, we are now studying the use of CBR

techniques in concert with others such as rule-based and model-based reasoning, that is, mixed paradigm systems.

3. CABARET — A Case-Based Reasoning Tool

In order to study issues concerning case-based reasoning, mixed paradigm systems integrating CBR with other modes of reasoning such as rule-based or model-based reasoning, and to facilitate the building of CBR systems, especially precedent-based (e.g., HYPO-style systems), our group has built an environment called CABARET (for CAsE-BAsed REasoning Tool).

CABARET is written in Common LISP and runs on both the MAC-II and the TI Explorer. It provides an integrated set of tools and facilities to support: (1) building precedent-based CBR systems; and (2) experimentation with multi-paradigm architectures. As a baseline for (1), we have required that CABARET support HYPO-like precedent-based systems, such as TAX-HYPO, which we described in the previous section. We are currently re-implementing, and extending, TAX-HYPO in CABARET.

3.1 General Considerations for a CBR Environment

As we have noted elsewhere [Rissland, 1987], the rudimentary components of any CBR system are the following:

1. **Case-Knowledge-Base (CKB)** — that is, a body of cases and experiences which is drawn upon for interpreting and/or solving the new, current case. For HYPO-like systems, the cases are highly structured objects and subobjects and they can be hypothetical (that is, not actually brought before or decided by some decision-maker) as well as real.
2. **Library of Indices** — that is, mechanisms to allow retrieval from the CKB. For HYPO-like systems, the indices, called "dimensions", are more than conjunctive combinations of Boolean features; they also contain information allowing assessment of how weak or strong a case is along the dimension by consideration of certain "focal" aspects [Rissland and Ashley, 1987; Rissland and Ashley, 1986]. For complex domains, like the law, the library of indices can have internal structure, such as a generalization/specialization hierarchy.
3. **Similarity/Relevancy Metrics** — that is, standards by which to assess the closeness of cases, judge their relevancy to the case at hand, and select "most on point" ("mope") cases. Such metrics allow the CBR system to cluster and rank cases, for both sides of an issue, and thus focus its attention on valuable cases. By attending to similarities and differences and to the pro/con impact of relevant cases, the reasoner is able to engage in fundamental CBR activities such as analogizing and distinguishing.
4. **Half-Order ¹²Theory of the Application Domain** — that is, hierarchies and taxonomies of knowledge specifically concerning the application domain. Such knowledge includes de-

¹²To borrow the term used in DENDRAL.

scriptors, variables/values, "factual predicates", and statutory predicates used to represent cases and indices. Of particular interest for dimension-like indices are partially ordered sets and symbolic hierarchies which are used as values for the "focal" slots. The rules or models constituting the rest of the domain theory are stored in the RBR "side" of CABARET.

5. **Justification/Explanation Methods** – that is, methods by which to justify the analysis and explain it. These are essential in precedent-based systems where one must reason according to the doctrine of precedent, *stare decisis*, which demands that the analysis of the new case be justified by appeal to past cases. Historically, these methods have not been found in problem-solving CBR systems, although there is no reason why they shouldn't be in future systems.

In addition for problem-solving CBR [Kolodner, 1987; Hammond, 1986], there must be the capability to modify and re-tailor solutions of past cases for the current case. Even though we speak no more of this aspect here, it is central and nontrivial one involving such techniques as adaptive planning.

For precedent-based CBR, there must also be:

1. **Precedent-Based Argumentation Capabilities** – that is, the capabilities to generate, and assess, precedent-based arguments. Knowledge necessary here includes knowledge of the status and ordering of cases according to court, decision maker, date, etc. For instance, a recent Supreme Court decision might be "worth" more than an older one; an Appeals Court decision, more than one from a District Court.
2. **Knowledge to Generate Hypotheticals** – that is, the capability to generate hypothetical cases to do such things as test the validity of an interpretation or argument by providing *gedanken* experiments as test cases or to flesh out a sparse CKB. One way to enable a CBR system to generate *useful* hypotheticals is to provide it with heuristics to guide it. Examples of such heuristics are *Enable a near-miss case* or *Generate an extreme case* [Rissland and Ashley, 1986].

3.2 CABARET in Support of Precedent-Based CBR

Our first goal is to have CABARET provide the needed support and functionality to enable one to build precedent-style CBR systems. Thus, it must provide tools for building, modifying, and maintaining CKB's, libraries of indices, various metrics of similarity, etc. described above.

In CABARET, we are not restricting ourselves to HYPO-like dimensional indices, although these typically require more knowledge and structure than most indices currently in use in the CBR community. On the other hand, we are also trying to support more structured memory than we used in HYPO, at the very least, CKB's that are networks of cases joined with typed links (e.g., in the law, such as "overruled" or "supports") and at some future time, perhaps, the sort of EMOP's memory employed by Kolodner's systems and their progeny. [Kolodner, 1983]. Also, CABARET is not restricted to legal applications.

Of course, issues about cases and indices go hand-in-hand and the design of CABARET has been made to reflect this, particularly, to reduce the burden on the system builder for explicitly building and maintaining obvious interconnections such as inverted lists of cases indexed by a given index or group of indices and the bookkeeping associated with adding and deleting cases from case memory.

Some specific CBR features of CABARET are the following:

- A three-tiered set of basic functions for cases and indices of: *definitions*, *prototypes*, and *instances*. The definitional level is used for defining objects, their components (e.g., slots) and their relations; this level is particularly useful when knowledge engineering a new application domain. The prototype level is for creating prototypical subclasses and instances of objects and relations. The instance level is used for building concrete instances of defined objects, especially during knowledge acquisition. Each of CABARET's standard toolkit functions has a version tuned to meet the requirements of each level and each type of object. Thus CABARET's basic functions "know" how to tailor themselves based on object type and tier without specification by the user.
- Basic functions for operating on cases and indices, such as CREATE, MODIFY, RENAME, DELETE, COPY, REVIVE and functions to SHOW, BROWSE, and VIEW individual objects, their components and groups of them (e.g., all the cases indexed by a set of indices).
- CABARET includes facilities to use or load different CKB's (e.g., only Federal cases, only cases recently entered), index libraries, and weighting schemes.
- Concerning weights, the default assumption is that all factors (e.g., indices) are treated equally, that is, given equal weight. However, CABARET does support other weighting assignments and allows the user to experiment with them. (This will probably prove interesting, if not important, for investigation of learning issues, although we are fully aware of the looming difficulties of the credit assignment problem. [Rissland and Ashley, 1988].)

In CABARET we are striving to provide ourselves with the sort of facilities to enable us to experiment with such things as change of similarity metrics and methods for computing most-on-point cases and neighborhoods of cases. For instance, HYPO used a measure of positive overlap of indices (dimensions) to compute similarity. There are many other possibilities such as symmetric difference. In HYPO, we did not weight features or dimensions and although we are not particularly interested in that approach, we realize others might be, so we have provided mechanisms to handle weightings of various kinds.

In HYPO, much of the domain-specific knowledge concerning the application was scattered about the implementation (e.g., dimensional mechanisms had their own store of knowledge about slot filler values, ranges, hierarchies, etc.). In CABARET, we are explicitly providing mechanisms to handle such half-order domain theories. We feel these explicit, independent mechanisms are particularly important in mixed paradigm systems where the other co-reasoner[s] (e.g., an expert system) must use the same domain knowledge as the case-based reasoner.

3.3 CABARET and Mixed Paradigm Systems

Our second goal is for CABARET's architecture to support systems that combine CBR and another type of inference module, such as a rule-based reasoner ("RBR") or a model-based reasoner ("MBR") [Goel, 1988; Koton, 1988]. Together, the CBR and its "co-reasoner" will work on precedent-type problems where one must reason with more than cases alone. As we mentioned, statutory domains, like tax, are examples where experts combine reasoning with rules and cases not only to find answers but also to justify them.

The control architecture we have chosen is agenda-based. Each co-reasoner (e.g., the CBR module and the RBR module) passes certain descriptors of the state of its internal processing to an agenda-handler which then decides which process to use and what goals to set for it based on such information, the current state of knowledge about the problem as a whole, and the goal of the reasoning. For instance, if the system is working on the rule-based "side" in a goal-driven backward chaining manner and the RBR halts before successfully establishing the goal, the failure is communicated to the agenda-handler which then might decide to suspend the rule-based reasoning and start up the CBR "side" with the goal of finding cases and arguments to support establishment of the unfulfilled goal. As an example, consider trying to establish a goal with five conjunctive subgoals. If the rule-based reasoner successfully established only four subgoals, the agenda-handler might direct the case-based reasoner to find cases that support the overall goal, yet were weak (or even completely deficient) on the missing subgoal. Such cases could be used to argue that these four out of five are good enough to establish the goal. This is a "disjunctify a conjunct" strategy; it uses cases to weaken the prerequisites of a rule to carve out an exception to it.

To implement CABARET's control strategies, we have had to ensure the agenda-handler access to data about each co-reasoner's inference state during processing. On the CBR side, we have followed HYPO's design, permitting direct access to internal data structures (e.g., Case Analysis Records, Claim Lattices and Argument Records). On the RBR side, we found ourselves unable to apply "as is" any rule-based shell available to us, and have built a generic rule-based module that permits access to intermediate processing data relevant to our control strategies (e.g., subgoals established/failed, nodes visited, state of the context, intermediate derived facts).

Such "glass-box" co-reasoners are monitored by demon processes, which translate observations of their designated co-reasoner into descriptors understandable to the agenda-handler. For instance, an RBR demon could recognize and pass on the observation that rule-based reasoning is stalled due to a failed fifth subgoal and that the four other conjunctive subgoals have been established — the RBR has a "near-miss" on a rule.

We realize that using such demonology and information-sharing between co-reasoners moves CABARET closer to traditional blackboard architectures. However, we do not anticipate in our applications of CABARET the multiplicities of levels of knowledge or knowledge sources present in typical blackboard applications, like speech recognition. We don't feel we have crossed the cost-benefit threshold into a realm where it is worthwhile to pay the overhead of maintaining a full-scale blackboard. Thus, we have not used a blackboard shell like Corkill's GBB (Generic BlackBoard development system) [Corkill, 1986], but reserve the use of GBB for possible further research.

The agenda-handler operates in the classic way: on the basis of observations and heuristic rules, it proposes, deletes, or changes the priority of tasks on an ordered agenda. As demonstrated in systems as Lenat's AM [Lenat, 1977], the mechanism is extremely elegant, powerful and suited to the kind of experiments we wish to do.

The following are examples of the sort of heuristics experts use to guide mixed paradigm (CBR and RBR) reasoning:

- *Fail \Rightarrow Switch*: If one mode of reasoning fails, switch to the other.
- *Rule-based Near Miss*: If all but one antecedent to a rule can be established, use CBR on the missing antecedent to establish it or to show the antecedent is not necessary.
- *All-fours*: As an initial attack on a problem, use CBR to determine if an extremely similar case is present in the CKB.
- *Drive with Rules*: As an initial attack on a problem, use RBR.
- *Driving with Lots of Facts*: If know lots of facts, try forward chaining.
- *Sanity Check*: Once a conclusion is reached, toggle the form of reasoning to determine if the same conclusion is derived. (Compare [Gardner, 1987].)
- *Deliberate Open Texture*: On deliberate open-textured predicates (e.g., "reasonable"), use CBR.
- *Cases Stacked for one Side*: If all the relevant cases are for one side or for one interpretation, try backward chaining on the stacked side's goal/interpretation.

3.4 Ms. Olivia's Tax Planning, Revisited

As a simplified example of how these heuristics might be applied in a mixed paradigm CBR/RBR system, let's reconsider the tax planning of our hypothetical musical taxpayer, Ms. Olivia. In particular, focus on her claim that her home office is her **PRINCIPAL PLACE OF BUSINESS**.

The law is in flux as to what test need be met to establish this fact. The current version of the rule of law used by the Tax Court is called the *focal point test*; it purports to situate the "focal point" — the principal place — of the taxpayer's business activities and determine whether the home office is the "focal point."

Probably because it simplifies evidentiary matters, this approach gives large weight to the traditional situs where an employee with a particular job would work. The test would find, for example, that the focal point of an airline pilot's job is in the cockpit, regardless of what other activities are required of the pilot (route planning, report generating, etc.), and where they are performed.

In a simplified rule form, the focal point test might be expressed:

If ((TAXPAYERS-EMPLOYMENT is a TRADE-OR-BUSINESS)
 and
 (SERVICES-VISIBLY-PROVIDED at HOME-OFFICE))
 then
 (HOME-OFFICE is PRINCIPAL-PLACE-OF-BUSINESS)

Ms. Olivia's home office would not qualify as her principal place of business under this rule. The first conjunct is satisfied since her activities do rise to the level of a trade or business, as required, and presumably this could be established using rule-based reasoning. However, the second conjunct is not satisfied: the focal point of her job, the place where her services are most visibly provided, is at Lincoln Center, not in her apartment. Under the applicable rule of law Ms. Olivia could not establish the second antecedent.

However, under these circumstances, a mixed paradigm reasoner's control would trigger the heuristic *Rule-based Near Miss*¹³, since the focal point rule fails on only one conjunct. With this direction, the CBR module would reason on the second conjunct. Given the case of Ms. Olivia, the CBR module of TAX-HYPO yields 3 most-on-point cases: Meiers, Weissman, and Drucker¹⁴ where the taxpayer won, receiving a deduction. In none of these cases was the SERVICES-VISIBLY-PROVIDED predicate satisfied. In each of these cases, a court took into consideration a different constellation of factors to establish the PRINCIPAL PLACE OF BUSINESS requirement. The state of the law on this point is not at a stage where a rule is recognized or can be stated, so resort to CBR is necessary to unearth these cases. The cases establish the argument that the second conjunct is really not a necessary condition for PRINCIPAL PLACE OF BUSINESS. Here, rule-based reasoning would fail to provide the desired result, whereas a mixed paradigm reasoner, driven by efficacious control heuristics, yields the solution.

4. Conclusions

In this paper we have:

1. Described the general problem of reasoning with cases in rule-governed domains, such as statutory legal domains like tax law and described the need for case-based interpretation of predicates in certain rule-governed applications.
2. Shown how our precedent-based CBR techniques developed for common law domains in the HYPO program were successfully applied in a statutory, rule-governed domain in the TAX-HYPO program, whose application domain is the home office deduction.

¹³or, possibly, *Fail* \Rightarrow *Switch*

¹⁴715 F.2d 67 (2d Cir. 1983) (on which Ms. Olivia's hypothetical case is based)

3. Described CABARET (CAsed BAsed REasoning Tool) whose two primary goals are to facilitate building precedent-based CBR applications, particularly of the HYPO variety and to provide an architecture to support mixed paradigm CBR systems.
4. Described some of the heuristics used in the agenda-based control architecture employed by CABARET and illustrated how they would work on an example from TAX-HYPO.

Based on our experiences, we are confident that the methodology of HYPO and TAX-HYPO are applicable to a variety of legal domains, including both common law and statutory domains. We hope that CABARET will help us build new applications of this type as well as new types of precedent-based reasoning involving CKB's, indexing methods and weighting schemes not employed in HYPO.

We are currently continuing our work on implementing CABARET and our research on questions regarding the use of agenda-based control mechanisms of CBR/RBR mixed paradigm reasoning systems.

REFERENCES

- | | |
|-----------------------------|--|
| [Ashley, 1988a] | Kevin D. Ashley. <i>Modelling Legal Argument: Reasoning with Cases and Hypotheticals</i> . Dept. of Computer and Information Science, University of Massachusetts (Doctoral Dissertation), 1988. |
| [Ashley and Rissland, 1988] | Kevin D. Ashley and Edwina L. Rissland. A Case-Based Approach to Modeling Legal Expertise. <i>IEEE Expert</i> , Fall 1988. |
| [Black, 1979] | Black's Law Dictionary. West Publishing Co., St. Paul, Minn., 5th Edition, 1979. |
| [BNA, 1986] | Kyler G. Knobbe. Tax Management Portfolio 241-4th (Hobby and Home Office Deductions — Sections 183 and 280A), Leonard L. Silverstein, Chief Editor, 1986. |
| [CCH, 1988] | Standard Federal Tax Reporter 1988. Commerce Clearing House, Inc., Chicago, 1988. |
| [Corkill, 1986] | Daniel D. Corkill, Kevin Q. Gallagher, and Kelly E. Murray. GBB: A Generic Blackboard Development System. In <i>Proceedings of the Fifth National Conference on Artificial Intelligence</i> , American Association for Artificial Intelligence, Philadelphia, Pa., 1986. |
| [Gardner, 1987] | Anne v.d.L. Gardner. <i>An Artificial Intelligence Approach to Legal Reasoning</i> , MIT Press, Cambridge, Mass., 1987. |

- [Goel, 1988] Ashok Goel and B. Chandrasekaran. Integrating Model Based Reasoning and Case-Based Reasoning for Design Problem Solving. In *Proceedings of the AAAI Workshop on AI in Design*. St. Paul, Minn., 1988.
- [Hammond, 1986] Kristian J. Hammond. CHEF: A Model of Case-Based Planning. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, American Association for Artificial Intelligence, Philadelphia, Pa., 1986.
- [Hart, 1961] H.L.A. Hart. *The Concept of Law*, Clarendon Press, Oxford, 1961.
- [Kolodner, 1987] Janet L. Kolodner. Extending Problem Solver Capabilities Through Case-Based Inference. In *Proceedings, Fourth Annual International Machine Learning Workshop*, Morgan Kaufman, Irvine, Calif., 1987.
- [Kolodner, 1983] Janet L. Kolodner. Maintaining Organization in a Dynamic Long-Term Memory. *Cognitive Science*, 7(4):243-280, 1983.
- [Koton, 1988] Phyllis Koton. Reasoning about Evidence in Causal Explanations. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, American Association for Artificial Intelligence, St. Paul, Minn., 1988.
- [Lakatos, 1976] Imre Lakatos. *Proofs and Refutations*. Cambridge University Press, Cambridge, 1976.
- [Lenat, 1977] Douglas B. Lenat. Automated Theory Formation in Mathematics. In *Proceedings of the Fifth Joint Conference on Artificial Intelligence*, International Joint Conference on Artificial Intelligence, Inc., Cambridge, Mass., 1977.
- [Prentice Hall, 1988] PII Federal Taxes 1988. Prentice-Hall, 1988.
- [Rissland, 1987] Edwina L. Rissland. *Research Initiative in Case-Based Reasoning*. Project Memo 17, The COUNSELOR Project, Department of Computer and Information Science, University of Massachusetts, 1987.
- [Rissland and Ashley, 1988] Edwina L. Rissland and Kevin D. Ashley. Credit Assignment and the Problem of Competing Factors in Case-Based Reasoning. In *Proceedings, DARPA Case-Based Reasoning Workshop*, Clearwater Beach, Fla., 1988.
- [Rissland and Ashley, 1987] Edwina L. Rissland and Kevin D. Ashley. A Case-Based System for Trade-Secrets Law. In *Proceedings, The First International Conference on Artificial Intelligence and Law*, Northeastern University, 1987.

- [Rissland and Ashley, 1986] Edwina L. Rissland and Kevin D. Ashley. Hypotheticals as Heuristic Device. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, American Association for Artificial Intelligence, Philadelphia, Pa., 1986.
- [Words and Phrases, 1964] *Words and Phrases*, Permanent Edition. West Publishing Co., St. Paul, Minn., 1964.